Synchronization-Aware and Algorithmically-Efficient Chance Constrained Optimal Power Flow

Russell Bent, Daniel Bienstock, Misha Chertkov

Columbia University, LANL

FERC Software Converence 2013



Review of past work: chance-constrained DC OPF

- CIGRE '09: large unexpected fluctuations in wind power can cause additional flows through the transmission system (grid)
- Large power deviations in renewables must be balanced by other sources, which may be far away
- Flow reversals may be observed control difficult
- A solution expand transmission capacity! Difficult (expensive), takes a long time
- Problems already observed when renewable penetration high

CIGRE -International Conference on Large High Voltage Electric Systems '09

- "Fluctuations" 15-minute timespan
- Due to turbulence ("storm cut-off")
- Variation of the same order of magnitude as mean
- Most problematic when renewable penetration starts to exceed 20 30%
- Many countries are getting into this regime



DC-OPF:

min c(p) (a quadratic)

s.t.

$$B\theta = p - d \tag{1}$$

$$|\beta_{ij}(\theta_i - \theta_j)| \le u_{ij}$$
 for each line ij (2)

$$P_g^{min} \leq p_g \leq P_g^{max}$$
 for each generator g (3)

Notation:

 $p = \text{vector of generations } \in \mathcal{R}^n, \quad d = \text{vector of loads } \in \mathcal{R}^n$ $B \in \mathcal{R}^{n \times n}, \quad \text{(bus susceptance matrix)}$



min
$$c(p)$$
 (a quadratic)

s.t.

$$B\theta = p - d$$

 $|\beta_{ij}(\theta_i - \theta_j)| \le u_{ij}$ for each line ij
 $P_g^{min} \le p_g \le P_g^{max}$ for each bus g

$$\begin{array}{ll} \min \ c(p) & \text{(a quadratic)} \\ \text{s.t.} & \\ B\theta = p - d \\ |\beta_{ij}(\theta_i - \theta_j)| \leq u_{ij} \quad \text{for each line } ij \\ P_g^{min} \leq p_g \leq P_g^{max} \quad \text{for each bus } g \end{array}$$

How does OPF handle short-term fluctuations in **demand** (d)? **Frequency control**:

- Automatic control: primary, secondary
- Generator output varies up or down proportionally to aggregate change

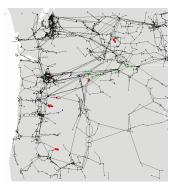
How does OPF handle short-term fluctuations in renewable output? **Answer:** Same mechanism, now used to handle aggregate wind power change



Experiment

Bonneville Power Administration data, Northwest US

- data on wind fluctuations at planned farms
- with standard OPF, 7 lines exceed limit $\geq 8\%$ of the time



Line trip model

summary: exceeding limit for too long is bad, but complicated want: "fraction time a line exceeds its limit is small" proxy: prob(violation on line i) $< \epsilon$ for each line i

Goals

- simple control
- aware of limits
- not too conservative
- computationally practicable

Control

For each generator i, two parameters:

- $\overline{p_i} = \text{mean output}$
- $\alpha_i = \text{response parameter}$

Real-time output of generator i:

$$p_i = \overline{p}_i - \alpha_i \sum_j \Delta \omega_j$$

where $\Delta \omega_j =$ change in output of renewable j (from mean).

$$\sum_{i} \alpha_{i} = 1$$

 \sim primary + secondary control



Computing line flows

wind power at bus i: $\mu_i + \mathbf{w}_i$

DC approximation

■
$$B\theta = \overline{p} - d$$

 $+(\mu + \mathbf{w} - \alpha \sum_{i \in G} \mathbf{w}_i)$

$$\bullet \theta = B^+(\bar{p} - d + \mu) + B^+(I - \alpha e^T)\mathbf{w}$$

flow is a linear combination of bus power injections:

$$\mathbf{f_{ij}} = \beta_{ij}(\boldsymbol{\theta}_i - \boldsymbol{\theta}_j)$$



Computing line flows

$$\mathbf{f}_{ij} = \beta_{ij} \left((B_i^+ - B_j^+)^T (\bar{p} - d + \mu) + (A_i - A_j)^T \mathbf{w} \right),$$
$$A = B^+ (I - \alpha e^T)$$

Given distribution of wind can calculate moments of line flows:

- $Ef_{ij} = \beta_{ij} (B_i^+ B_j^+)^T (\bar{p} d + \mu)$
- $var(\mathbf{f_{ij}}) := s_{ij}^2 \ge \beta_{ij}^2 \sum_k (A_{ik} A_{jk})^2 \sigma_k^2$ (assuming independence)
- and higher moments if necessary



Chance constraints to deterministic constraints

- lacktriangledown chance constraint: $P(\mathbf{f_{ij}} > f_{ij}^{max}) < \epsilon_{ij}$ and $P(\mathbf{f_{ij}} < -f_{ij}^{max}) < \epsilon_{ij}$
- from moments of f_{ij} , can get conservative approximations using e.g. Chebyshev's inequality

Chance constraints to deterministic constraints

- lacktriangledown chance constraint: $P(\mathbf{f_{ij}} > f_{ij}^{max}) < \epsilon_{ij}$ and $P(\mathbf{f_{ij}} < -f_{ij}^{max}) < \epsilon_{ij}$
- lacktriangleright from moments of f_{ij} , can get conservative approximations using e.g. Chebyshev's inequality
- \blacksquare for Gaussian wind, can do better, since f_{ij} is Gaussian :

$$|E\mathbf{f}_{ij}| + var(\mathbf{f}_{ij})\phi^{-1}(1 - \epsilon_{ij}) \le f_{ij}^{max}$$



Formulation:

Choose mean generator outputs and control to minimize expected cost, with the probability of line overloads kept small.

$$\begin{split} & \min_{\overline{p},\alpha} \mathbb{E}[c(\overline{p})] \\ \text{s.t.} & \sum_{i \in G} \alpha_i = 1, \ \alpha \geq 0 \\ & B\delta = \alpha, \delta_n = 0 \\ & \sum_{i \in G} \overline{p}_i + \sum_{i \in W} \mu_i = \sum_{i \in D} d_i \\ & \overline{f}_{ij} = \beta_{ij} (\overline{\theta}_i - \overline{\theta}_j), \\ & B\overline{\theta} = \overline{p} + \mu - d, \ \overline{\theta}_n = 0 \\ & s_{ij}^2 \geq \beta_{ij}^2 \sum_{k \in W} \sigma_k^2 (B_{ik}^+ - B_{jk}^+ - \delta_i + \delta_j)^2 \\ & |\overline{f}_{ij}| + s_{ij}\phi^{-1} (1 - \epsilon_{ij}) \leq f_{ij}^{max} \end{split}$$

Formulation:

Choose mean generator outputs and control to minimize expected cost, with the probability of line overloads kept small.

$$\begin{split} \min_{\overline{p},\alpha} \mathbb{E}[c(\overline{p})] \\ \text{s.t.} \sum_{i \in G} \alpha_i &= 1, \ \alpha \geq 0 \\ B\delta &= \alpha, \delta_n = 0 \\ \sum_{i \in G} \overline{p}_i + \sum_{i \in W} \mu_i &= \sum_{i \in D} d_i \\ \overline{f}_{ij} &= \beta_{ij} (\overline{\theta}_i - \overline{\theta}_j), \\ B\overline{\theta} &= \overline{p} + \mu - d, \ \overline{\theta}_n = 0 \\ s_{ij}^2 &\geq \beta_{ij}^2 \sum_{k \in W} \sigma_k^2 (B_{ik}^+ - B_{jk}^+ - \delta_i + \delta_j)^2 \\ |\overline{f}_{ij}| &+ s_{ij} \phi^{-1} (1 - \epsilon_{ij}) \leq f_{ii}^{max} \end{split}$$

A convex optimization problem.



Big cases

Polish 2003-2004 winter peak case

- 2746 buses, 3514 branches, 8 wind sources
- 5% penetration and $\sigma = .3\mu$ each source

CPLEX: the optimization problem has

- 36625 variables
- 38507 constraints, 6242 conic constraints
- 128538 nonzeros, 87 dense columns



Big cases

Polish 2003-2004 winter peak case

- 2746 buses, 3514 branches, 8 wind sources
- 5% penetration and $\sigma = .3\mu$ each source

CPLEX: the optimization problem has

- 36625 variables
- 38507 constraints, 6242 conic constraints
- 128538 nonzeros, 87 dense columns



Big cases

CPLEX:

- total time on 16 threads = 3393 seconds
- "optimization status 6"
- solution is wildly infeasible

Gurobi:

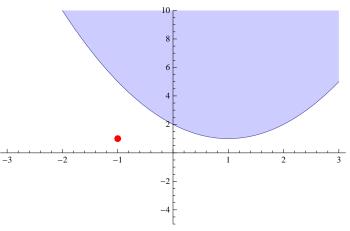
- time: 31.1 seconds
- "Numerical trouble encountered"

overview

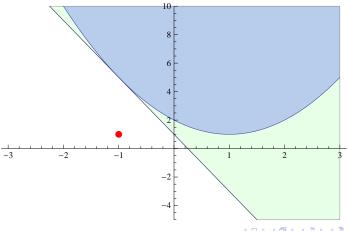
Cutting-plane algorithm:

```
remove all conic constraints
repeat until convergence:
    solve linearly constrained problem
    if no conic constraints violated: return
    find separating hyperplane for maximum violation
    add linear constraint to problem
```

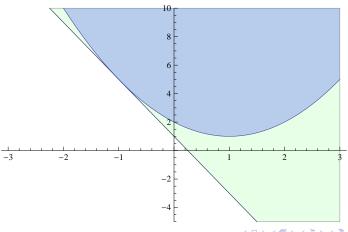
Candidate solution violates conic constraint



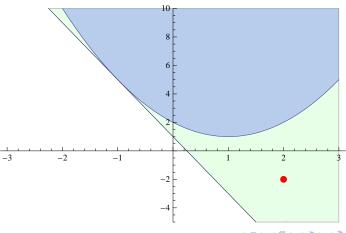
Separate: find a linear constraint also violated



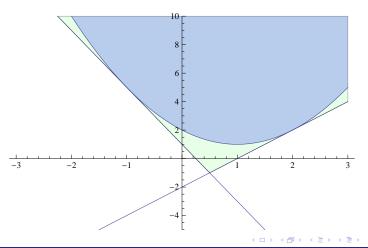
Solve again with linear constraint



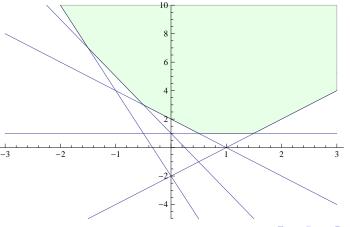
New solution still violates conic constraint



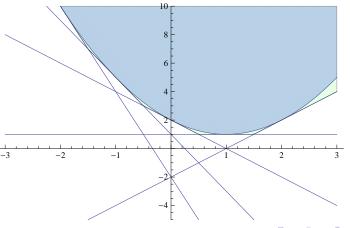
Separate again



We might end up with many linear constraints

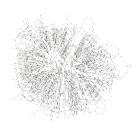


... which approximate the conic constraint



Polish 2003-2004 case CPLEX: "opt status 6"

Gurobi: "numerical trouble"



Example run of cutting-plane algorithm:

Iteration	Max rel. error	Objective
1	1.2e-1	7.0933e6
4	1.3e-3	7.0934e6
7	1.9e-3	7.0934e6
10	1.0e-4	7.0964e6
12	8.9e-7	7.0965e6

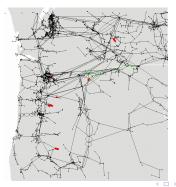
Total running time: 32.9 seconds



Back to motivating example

BPA case

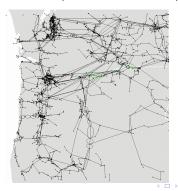
- standard OPF: cost 235603, 7 lines unsafe ≥ 8% of the time
- CC-OPF: cost 237297, every line safe ≥ 98% of the time
- run time = 9.5 seconds (one cutting plane!)



Back to motivating example

BPA case

- standard OPF: cost 235603, 7 lines unsafe ≥ 8% of the time
- CC-OPF: cost 237297, every line safe ≥ 98% of the time
- run time = 9.5 seconds (one cutting plane!)



Summary: Bienstock, Chertkov, Harnett 2012

- Specialized cutting-plane algorithm proves effective
- Commercial solvers do not
- Algorithm efficient even in cases with thousands of buses/lines
- Algorithm can be made robust with respect to data errors

Summary: Bienstock, Chertkov, Harnett 2012

- Specialized cutting-plane algorithm proves effective
- Commercial solvers do not
- Algorithm efficient even in cases with thousands of buses/lines
- Algorithm can be made robust with respect to data errors

Can we handle power flows more accurately?



Active power, lossless OPF:

$$\min_{p,\theta} c(p)$$

s.t.

$$\sum_{i:i\in\mathcal{L}}\beta_{ij}\sin(\theta_i-\theta_j) = p_i - d_i \quad \forall i\in\mathcal{B}$$
 (4)

$$|\beta_{ij}\sin(\theta_i-\theta_j)| \le u_{ij}$$
 for each line ij (5)

$$P_g^{min} \leq p_g \leq P_g^{max}$$
 for each generator g (6)

From Boyd (2012): Suppose you solve the convex optimization problem:

$$\min \ \sum_{ij \in \mathcal{L}} \beta_{ij} \Psi(\rho_{ij})$$

s.t.

$$\sum_{j:ij\in\mathcal{L}}\beta_{ij}\rho_{ij} - \sum_{j:ji\in\mathcal{L}}\beta_{ij}\rho_{ji} = p_i - d_i \quad \forall i \in \mathcal{B}$$
 (7)

$$|\rho_{ij}| < \min\{1, u_{ij}/\beta_{ij}\}$$
 for each line ij (8)

From Boyd (2012): Suppose you solve the convex optimization problem:

$$\min \ \sum_{ij\in\mathcal{L}} \beta_{ij} \Psi(\rho_{ij})$$

s.t.

$$\sum_{j:ij\in\mathcal{L}}\beta_{ij}\rho_{ij} - \sum_{j:ji\in\mathcal{L}}\beta_{ij}\rho_{ji} = p_i - d_i \quad \forall i\in\mathcal{B}$$
 (7)

$$|\rho_{ij}| < \min\{1, u_{ij}/\beta_{ij}\}$$
 for each line ij (8)

Where for $|\rho| < 1$,

$$\Psi(\rho) \doteq \int_{-1}^{\rho} \arcsin(y) \, dy$$

From Boyd (2012): Suppose you solve the convex optimization problem:

$$\min \ \sum_{ij \in \mathcal{L}} \beta_{ij} \Psi(\rho_{ij})$$

s.t.

$$\sum_{j:ij\in\mathcal{L}}\beta_{ij}\rho_{ij} - \sum_{j:ji\in\mathcal{L}}\beta_{ij}\rho_{ji} = p_i - d_i \quad \forall i\in\mathcal{B}$$
 (7)

$$|\rho_{ij}| < \min\{1, u_{ij}/\beta_{ij}\}$$
 for each line ij (8)

Where for $|\rho| < 1$,

$$\Psi(\rho) \doteq \int_{-1}^{\rho} \arcsin(y) \, dy$$

Then: If θ_i is the optimal dual for (7), $\rho_{ij} = \sin(\theta_i - \theta_j)$.

From Boyd (2012): Suppose you solve the convex optimization problem:

$$\min \ \sum_{ij \in \mathcal{L}} \beta_{ij} \Psi(\rho_{ij})$$

s.t.

$$\sum_{j:ij\in\mathcal{L}}\beta_{ij}\rho_{ij} - \sum_{j:ji\in\mathcal{L}}\beta_{ij}\rho_{ji} = p_i - d_i \quad \forall i\in\mathcal{B}$$
 (7)

$$|\rho_{ij}| < \min\{1, u_{ij}/\beta_{ij}\}$$
 for each line ij (8)

Where for $|\rho| < 1$,

$$\Psi(\rho) \doteq \int_{-1}^{\rho} \arcsin(y) \, dy$$

Then: If θ_i is the optimal dual for (7), $\rho_{ij} = \sin(\theta_i - \theta_j)$.

How can we incorporate this methodology into OPF-type problems?

Suppose you solve the **convex optimization problem**:

$$\min_{p,\rho,\delta\geq 0} c(p) + D \sum_{ij\in\mathcal{L}} \beta_{ij} \Psi(\rho_{ij}) - K \sum_{ij\in\mathcal{L}} \beta_{ij} \log(\delta_{ij})$$
 (9)

s.t.

$$\sum_{j:ij\in\mathcal{L}}\beta_{ij}\rho_{ij} - \sum_{j:ji\in\mathcal{L}}\beta_{ij}\rho_{ji} = p_i - d_i \quad \forall i\in\mathcal{B}$$
 (10)

$$|
ho_{ij}| + \min\{1, u_{ij}/eta_{ij}\}\delta_{ij} < \min\{1, u_{ij}/eta_{ij}\}$$
 for each line **(11)**

$$P_{g}^{min} \leq p_{g} \leq P_{g}^{max}$$
 for each generator g

For appropriate positive constants **D** (small) and **K** (large).

Suppose you solve the **convex optimization problem**:

$$\min_{p,\rho,\delta\geq 0} c(p) + D \sum_{ij\in\mathcal{L}} \beta_{ij} \Psi(\rho_{ij}) - K \sum_{ij\in\mathcal{L}} \beta_{ij} \log(\delta_{ij})$$
 (9)

s.t.

$$\sum_{j:ij\in\mathcal{L}}\beta_{ij}\rho_{ij} - \sum_{j:ji\in\mathcal{L}}\beta_{ij}\rho_{ji} = p_i - d_i \quad \forall i\in\mathcal{B}$$
 (10)

$$|
ho_{ij}| + \min\{1, u_{ij}/eta_{ij}\}\delta_{ij} < \min\{1, u_{ij}/eta_{ij}\}$$
 for each line **(11)**

$$P_g^{min} \leq p_g \leq P_g^{max}$$
 for each generator g

For appropriate positive constants D (small) and K (large). Theorem:

(1) The optimal ρ_{ij} are approximate optimal active flows

Suppose you solve the **convex optimization problem**:

$$\min_{p,\rho,\delta\geq 0} c(p) + D \sum_{ij\in\mathcal{L}} \beta_{ij} \Psi(\rho_{ij}) - K \sum_{ij\in\mathcal{L}} \beta_{ij} \log(\delta_{ij})$$
 (9)

s.t.

$$\sum_{j:ij\in\mathcal{L}}\beta_{ij}\rho_{ij} - \sum_{j:ji\in\mathcal{L}}\beta_{ij}\rho_{ji} = p_i - d_i \quad \forall i\in\mathcal{B}$$
 (10)

$$|
ho_{ij}| + \min\{1, u_{ij}/eta_{ij}\}\delta_{ij} < \min\{1, u_{ij}/eta_{ij}\}$$
 for each line **(11)**

$$P_g^{min} \leq p_g \leq P_g^{max}$$
 for each generator g

For appropriate positive constants D (small) and K (large). Theorem:

- (1) The optimal ρ_{ii} are approximate optimal active flows
- (2) $\rho_{ij} \approx \sin(\theta_i \theta_i)$ $\theta = \text{optimal duals to (10)}.$



Somewhat more general: $\gamma_{ij}=$ sine of max phase difference on ij

$$\min_{p,\rho,\delta\geq 0} c(p) + D \sum_{ij\in\mathcal{L}} \beta_{ij} \Psi(\rho_{ij}) - K \sum_{ij\in\mathcal{L}} \beta_{ij} \log(\delta_{ij})$$
 (12)

s.t.

$$\sum_{j:ij\in\mathcal{L}}\beta_{ij}\rho_{ij} - \sum_{j:ji\in\mathcal{L}}\beta_{ij}\rho_{ji} = p_i - d_i \quad \forall i \in \mathcal{B}$$
 (13)

$$|
ho_{ij}| + \min\{\gamma_{ij}, u_{ij}/eta_{ij}\}\delta_{ij} < \min\{\gamma_{ij}, u_{ij}/eta_{ij}\}$$
 for each line (1)4)

$$P_g^{min} \leq p_g \leq P_g^{max}$$
 for each generator g

For appropriate positive constants D (small) and K (large). Theorem:

- (1) The optimal ρ_{ij} are approximate optimal active flows
- (2) $\rho_{ij} \approx \sin(\theta_i \theta_i)$ $\theta = \text{optimal duals to (13)}.$

Ongoing work:

$$\begin{split} \min_{\rho,\,\rho,\,\delta\geq 0} \ c(\rho) \ + \ D \sum_{ij\in\mathcal{L}} \beta_{ij} \Psi(\rho_{ij}) \ - \ K \sum_{ij\in\mathcal{L}} \beta_{ij} \log(\delta_{ij}) \end{split}$$
 s.t.
$$\sum_{j:ij\in\mathcal{L}} \beta_{ij} \rho_{ij} \ - \sum_{j:ji\in\mathcal{L}} \beta_{ij} \rho_{ji} \ = \ \rho_i \ - \ d_i \qquad \forall i\in\mathcal{B}$$

$$|\rho_{ij}| \ + \ \min\{1,u_{ij}/\beta_{ij}\}\delta_{ij} \ < \ \min\{1,u_{ij}/\beta_{ij}\} \quad \text{for each line } ij$$

$$P_g^{min} \ \leq \ \rho_g \ \leq \ P_g^{max} \quad \text{for each generator } g \end{split}$$

- $lue{}$ Outer envelope approximation to functions $lue{}$ $lue{}$ lue
- $D \rightarrow 0$, $K \rightarrow +\infty$ needs to be managed
- Existing methodology for logarithmic barrier algorithms can be leveraged
- Early infeasibility detection can be important



Dörfler, Chertkov, Bullo 2013: an approximation

$$\begin{array}{ll} \displaystyle \min_{p,\vartheta} \ c(p) \\ \text{s.t.} \\ \displaystyle \sum_{j:ij\in\mathcal{L}} \beta_{ij} (\vartheta_i - \vartheta_j) \ = \ p_i - d_i \qquad \forall i \in \mathcal{B} \\ \\ |\vartheta_i - \vartheta_j| < \ \min\{1, u_{ij}/\beta_{ij}\} \quad \text{for each line } ij \end{array}$$

- lacktriangle The artheta are auxiliary variables only
- In experiments, $\vartheta_i \vartheta_j$ provides a close approximation to the lossless (active) AC power flow on each line ij
- (But does not provide phase angles)



A combination of two ideas

■ On any line ij, we replace $\sin(\theta_i - \theta_j)$ with the quantity $\vartheta_i - \vartheta_j$

A combination of two ideas

- lacksquare On any line ij, we replace $\sin(heta_i- heta_j)$ with the quantity $heta_i- heta_j$
- So 'sync' constraint $|\sin(\theta_i \theta_j)| \le \gamma_{ij}$ becomes $|\vartheta_i \vartheta_j| \le \gamma_{ij}$

A combination of two ideas

- On any line ij, we replace $\sin(\theta_i \theta_j)$ with the quantity $\vartheta_i \vartheta_j$
- So 'sync' constraint $|\sin(\theta_i \theta_j)| \le \gamma_{ij}$ becomes $|\vartheta_i \vartheta_j| \le \gamma_{ij}$
- But in either case the constraint is stochastic

A combination of two ideas

- lacksquare On any line ij, we replace $\sin(heta_i- heta_j)$ with the quantity $heta_i- heta_j$
- So 'sync' constraint $|\sin(\theta_i \theta_j)| \le \gamma_{ij}$ becomes $|\vartheta_i \vartheta_j| \le \gamma_{ij}$
- But in either case the constraint is stochastic

Chance-constrained version: $P(|\vartheta_i - \vartheta_j| > \gamma_{ij}) < \epsilon_{ij}$

Example: $\epsilon_{ij} = 10^{-4}$.



Control (again)

For each generator i, two parameters:

- $\overline{p_i}$ = mean output
- $\alpha_i = \text{response parameter}$

Real-time output of generator i:

$$p_i = \overline{p}_i - \alpha_i \sum_i \Delta \omega_j$$

where $\Delta\omega_j =$ change in output of renewable j (from mean).

$$\sum_{i} \alpha_{i} = 1$$



Control (again)

For each generator i, two parameters:

- $\overline{p_i}$ = mean output
- $\alpha_i = \text{response parameter}$

Real-time output of generator i:

$$p_i = \overline{p}_i - \alpha_i \sum_i \Delta \omega_j$$

where $\Delta\omega_i$ = change in output of renewable j (from mean).

$$\sum_{i} \alpha_{i} = 1$$

So for any line ij, $\vartheta_{\bf i} - \vartheta_{\bf j} = \sum_k a_k (\bar p_k - d_k + \mu_k) + \sum_k b_k \omega_{\bf k}$



Chance-constrained, thermal and sync-aware (approximate) OPF:

Choose mean generator outputs and control to minimize expected cost, with the probability of line overloads and phase angle excursions kept small. (abridged)

$$\begin{split} & \min_{\overline{p},\alpha} \mathbb{E}[c(\overline{p})] \\ \text{s.t.} & \sum_{i \in G} \alpha_i = 1, \ \alpha \geq 0 \\ & B\delta = \alpha \\ & \sum_{ij \in \mathcal{L}} \beta_{ij} (\overline{\vartheta}_i - \overline{\vartheta}_j) = \overline{p_i} + \mu_{\mathbf{i}} - d_i \\ & P(\beta_{ij} | \vartheta_{\mathbf{i}} - \vartheta_{\mathbf{j}} | > u_{ij}) \leq \epsilon_1 \quad \text{for each line } ij \\ & P(|\vartheta_{\mathbf{i}} - \vartheta_{\mathbf{j}}| > \gamma_{ij}) \leq \epsilon_2 \quad \text{for each line } ij \\ & P(\mathbf{p_g} < P_g^{min} \text{ or } P_g^{max} < \mathbf{p_g}) \leq \epsilon_3 \quad \text{for each generator } g \end{split}$$

$$\epsilon_2 \ll \epsilon_3 \ll \epsilon_1$$

Again: a conic optimization problem



Summary of computational experiments

- On Polish grid example (approximately 3000 buses, 388 generators and 3799 lines), cutting-plane algorithm converges within 5-30 seconds and 2-30 iterations on a current computer
- Algorithm 'discovers' at-risk lines
- Fairly smooth convergence with decreased risk as the generation dispatch and control parameters are improved
- Geographical patterns of at-risk lines exposed
- Standard OPF produces poor solutions risky and expensive
- See paper!

